DiGiTAL TOTAL

Beitrag ID: **62** Beitragskennung: **57**   Typ: **Poster**

INEL –Grammar, Corpora and Language Technology for Indigenous Northern Eurasian Languages

Ever since 2016, the spanning 18 years long-term project INEL has been generating deeply annotated language corpora and accompanying digital resources using existing and acquiring new language material from a number of heavily endangered languages of the Northern Eurasian Area. The core aim is not only to make these resources sustainably available after their publication but also provide continuous data curation and analysis already during their preparation. This puts high demands on the digital workflows that involve various data formats, tools, and approaches at all levels from data preprocessing to the publication of the corpus. The tools range from well-established software for linguistic work to frameworks and tools developed and maintained specifically for the project needs.

Find me @ my poster

2,3

Keywords

language corpora,
workflow management,
quality assurance,
data sustainability

Autoren: LAZARENKO, Elena (Universität Hamburg);   Herr RIAPOSOV, Aleksandr (Universität Hamburg)